

## The distribution and predictive value of Bishop scores in nulliparas between 37 and 42 weeks gestation

PETER E. NIELSEN, BOBBY C. HOWARD, TAMI CRABTREE, ALISON L. BATIG, & JASON A. PATES

Madigan Health System, Tacoma, Washington, USA

### Abstract

**Objective.** The natural distribution and predictive accuracy of Bishop scores was evaluated to predict cesarean delivery (CD) in nulliparas between 37 and 42 weeks gestation.

**Study design.** Subjects underwent serial digital cervical examinations. The Bishop score was evaluated as a binary and continuous factor to predict CD at each gestational week beginning at 37 weeks. Bishop scores were categorized as  $\leq 5$  or  $> 5$ , and CD rates were compared across Bishop score categories using chi square or Fisher exact tests at each gestational week beginning at 37 weeks.

**Results.** In all, 171 patients were prospectively followed. The overall CD rate was 27.5%. The prevalence of unfavorable Bishop scores, categorized as  $\leq 5$ , decreased with increasing gestation age until 41 weeks. CD rates for the cohort with unfavorable Bishop scores was higher than those with favorable scores at each week. The likelihood ratio for CD was 1.35-2.00, depending on gestational age. The Bishop score that best predicted subsequent vaginal delivery following expectant management was  $> 3$  at 37 weeks and  $> 5$  at 39 weeks.

**Conclusion.** A Bishop score  $\leq 5$  between 37 and 39 weeks gestation predicts a higher CD rate compared to patients with a Bishop score  $> 5$  implying an intrinsically higher CD risk despite expectant management.

**Keywords:** Cesarean, Bishop score, distribution

### Introduction

Bishop scores have been used to predict the success of labor induction at term [1]. In addition, vaginal delivery rates following induction are similar to spontaneous labor when the preinduction Bishop score exceeds 8 in nulliparas [2]. Retrospective studies evaluating labor induction on nulliparas with an unfavorable cervix (Bishop score of 8 or less) demonstrate a two fold higher cesarean delivery (CD) rate in those who are induced. [3-9]. However, prospective randomized trials evaluating induction of labor suggest either no increased risk [10-13] or a lower risk of CD [14]. Unfortunately, the observational studies compared induction with spontaneous labor instead of expectant management. The only option for providers or patients is either induction or expectant management, not spontaneous labor, since neither the patient nor the provider can ensure this outcome [15].

Evaluation of potential factors for predicting successful vaginal delivery at term is important since maternal fetal complications increase at gestational ages beyond 38 weeks [16]. In addition, characterizing Bishop scores and expectant management delivery outcomes in term nulliparous patients may guide management options. Finally, these data could help design a study evaluating the effect of elective induction in nulliparas, since the rate of CD in patients expectantly

managed is the appropriate control group rate and avoids selection bias associated with comparing electively induced patients with those who present in spontaneous labor.

The purpose of this investigation was to evaluate the distribution of Bishop scores, the CD rate based on expectant management by Bishop score and gestational age, and the ability to predict CD in nulliparas between 37 and 42 weeks gestation.

### Materials and methods

This was an Institutional Review Board (IRB) approved prospective observational study. Potential study participants were identified between 34 and 37 weeks gestation. Study enrollment began in May 2005 and continued through the proscribed cutoff of March 2007. Individuals who met the following inclusion criteria were offered enrollment: no known indication for scheduled delivery, nulliparous, singleton gestation, cephalic presentation, age 18 to 40 years, no known fetal anomaly, no contraindication to labor or vaginal delivery, intact amniotic membranes, reliable for follow up. After informed consent, enrolled subjects were examined at each subsequent routine clinic visit beginning not earlier than 37 weeks gestation. A standardized Bishop scoring data sheet was used and completed immediately following the examination. Examinations

(Received 21 October 2010; revised 16 March 2011; accepted 17 March 2011)

Correspondence: Peter E. Nielsen, Department of Obstetrics and Gynecology, Madigan Army Medical Center, MCHJ-OG (ATTN: COL Nielsen), Tacoma, WA 98431, USA. Tel: +253-968-5161. Fax: +253-968-5508. E-mail: Peter.Nielsen@amedd.army.mil

Report Documentation Page			Form Approved OMB No. 0704-0188	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE <b>2011</b>	2. REPORT TYPE	3. DATES COVERED <b>00-00-2011 to 00-00-2011</b>		
<b>4. TITLE AND SUBTITLE</b> <b>The Distribution And Predictive Value Of Bishop Scores In Nulliparas Between 37 And 42 Weeks Gestation</b>			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
<b>6. AUTHOR(S)</b>			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> <b>Madigan Army Medical Center, MCHJ-OG (ATTN: COL Nielsen, Department of Obstetrics and Gynecology, Tacoma, WA, 98431</b>			8. PERFORMING ORGANIZATION REPORT NUMBER	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> <b>Approved for public release; distribution unlimited</b>				
<b>13. SUPPLEMENTARY NOTES</b> <b>Journal of Maternal-Fetal and Neonatal Medicine, Early Online, pgs 1-5, 2011</b>				
<b>14. ABSTRACT</b> <p><b>Objective.</b> The natural distribution and predictive accuracy of Bishop scores was evaluated to predict cesarean delivery (CD) in nulliparas between 37 and 42 weeks gestation. <b>Study design.</b> Subjects underwent serial digital cervical examinations. The Bishop score was evaluated as a binary and continuous factor to predict CD at each gestational week beginning at 37 weeks. Bishop scores were categorized as 5 or 45, and CD rates were compared across Bishop score categories using chi-square or Fisher exact tests at each gestational week beginning at 37 weeks. <b>Results.</b> In all, 171 patients were prospectively followed. The overall CD rate was 27.5%. The prevalence of unfavorable Bishop scores, categorized as 5, decreased with increasing gestation age until 41 weeks. CD rates for the cohort with unfavorable Bishop scores was higher than those with favorable scores at each week. The likelihood ratio for CD was 1.35?2.00, depending on gestational age. The Bishop score that best predicted subsequent vaginal delivery following expectant management was 43 at 37 weeks and 45 at 39 weeks. <b>Conclusion.</b> A Bishop score 5 between 37 and 39 weeks gestation predicts a higher CD rate compared to patients with a Bishop score 45 implying an intrinsically higher CD risk despite expectant management.</p>				
<b>15. SUBJECT TERMS</b>				
<b>16. SECURITY CLASSIFICATION OF:</b>  a. REPORT <b>unclassified</b>			<b>17. LIMITATION OF ABSTRACT</b>  Same as Report (SAR)	<b>18. NUMBER OF PAGES</b>  <b>6</b>
b. ABSTRACT <b>unclassified</b> c. THIS PAGE <b>unclassified</b>				
<b>19a. NAME OF RESPONSIBLE PERSON</b>				

were repeated at each visit until delivery. The Bishop score was calculated at each examination using the cervical dilation, effacement, consistency, position, and station [1]. The cervical assessments were performed by the particular staff or resident obstetrician conducting the office visit with the patient and the previous Bishop score was not available to this examiner. Delivery data were collected for each participant and categorized as either a vaginal or CD. Deliveries that occurred vaginally with use of forceps or vacuum assistance were categorized as vaginal deliveries. Expectant management was defined as weekly routine obstetric visits and delivery planned only for the development of maternal or fetal indications or by 42 weeks gestation.

Using a logistic regression to determine the odds ratio (OR), the Bishop score was evaluated as both a binary and a continuous factor to predict CD at each gestational week beginning at 37 weeks. For the binary evaluation, Bishop scores were categorized as  $\leq 5$  or  $> 5$  as this Bishop score cutoff was previously studied in our institution and found to be clinically useful in distinguishing a favorable from an unfavorable cervix [13]. Using this binary categorization, CD rates were compared across Bishop score categories using the chi square and Fisher exact test at each gestational week beginning at 37 weeks. The Bishop score at a given observational window was used to evaluate subsequent CD at any time following this window. The sensitivity, specificity, positive and negative likelihood ratios, and positive predictive value (PPV) and negative predictive value (NPV) of the Bishop score was determined. A receiver operator characteristic (ROC) curve was also created to further evaluate Bishop score as predictor of CD across multiple potential cut off values. The *p* values were calculated to test the null hypothesis that the area under the ROC curve was less than or equal to 0.5 (indicating a useless predictor). For all hypothesis tests, a *p* value of less than 0.05 was considered statistically significant. A priori sample size calculation for this study was not performed as the actual incidence of Bishop scores and degree of change across a range of gestational ages was unknown at the time of protocol preparation.

## Results

Three hundred and seventeen patients were identified for possible enrollment between 34 and 37 weeks gestation. Of these, 233 rendered their consent for participation and 84 declined enrollment. Following enrollment, 62 additional patients were excluded. Reasons for exclusion included withdrawal of consent, failure to present for follow up beyond 37 weeks, and delivery prior to 37 weeks. Subjects were also withdrawn from the evaluation if they did not meet inclusion criterion following enrollment but before 37 gestational weeks (i.e., developed an indication for induction or primary CD after enrollment but before 37 weeks gestation). There were 171 evaluable patients who delivered after 37 gestational weeks and for whom cervical examination and delivery data were available at the completion of the study (Figure 1).

The mean maternal age was 23.6 years (SD = 4.1) and the mean gestational age at enrollment was 37.6 weeks (SD = 0.6). Eighty patients had their first Bishop score assessment at 37 weeks gestation and 91 had their first Bishop score assessment at 38 weeks gestation. Thirty five percent (59/171) of patients required induction of labor after enrollment due to the development of a maternal or fetal indication. Specifically, 69% of inductions were for postdatism (41 weeks or greater),

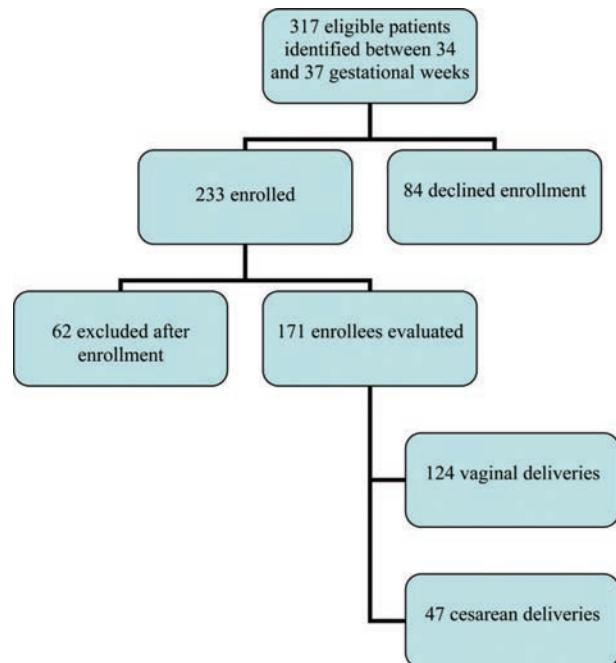


Figure 1. Study enrollment.

19% were for pregnancy induced hypertension, 7% for premature rupture of membranes, and the remaining 5% (1 patient each) for unremitting backpain, oligohydramnios, and nonreassuring fetal heart tracing. The overall CD rate was 27.5% (47/171). The CD rate of those patients who required induction of labor compared to those who presented in spontaneous labor was 42.4% (25/59) versus 19.6% (22/112). The CD rate at each gestational week was 25.0% (2/8) at 37 week, 25.9% (7/27) at 38 weeks, 31.1% (19/61) at 39 weeks, 19.6% (9/46) at 40 weeks, and 38.5% (10/26) at 41 weeks. Only one patient in the study delivered at 42 weeks gestation and this delivery occurred vaginally. All patients included in this evaluation delivered between 37 and 42 gestational weeks, with the highest percentage (36.1%) delivering at 39 weeks. Ninety one percent of patients had an ultrasound examination performed at  $\leq 20$  weeks gestation and 52% had an ultrasound examination performed at less than 13 weeks gestation confirming gestational age. Last menstrual period dating and an ultrasound examination performed at greater than 20 weeks gestation were used to assess gestational age in the remaining 9% of patients.

ROC curves evaluating the relationship between Bishop scores by gestational age and subsequent CD were created for each gestational week from 37 to 40. The ROC curves at 37 and 39 weeks demonstrated the largest area under the curve at 0.77 and 0.73, respectively (Figures 2 and 3), and both were statistically significant (*p* < 0.0001). The Bishop score that best predicted subsequent vaginal delivery following expectant management was  $> 3$  at 37 weeks and  $> 5$  at 39 weeks. The sensitivity of a Bishop score of 3 at 37 weeks was 80% and for a Bishop score of 5 at 39 weeks was 73%.

The prevalence of unfavorable Bishop scores, defined as those scores  $\leq 5$ , decreased with increasing gestation age until 41 gestational weeks. At 37 gestational weeks, 71.3% [57/80, 95% confidence interval (CI): 60.5 80.0] of enrollees had a Bishop score  $\leq 5$ . The rate was 63.7% (79/124 95% CI: 55.0 71.6) at 38 gestational weeks, 46.6% (48/103, 95% CI: 37.3 56.2) at 39 gestational weeks, 40.4% (23/57, 95% CI:

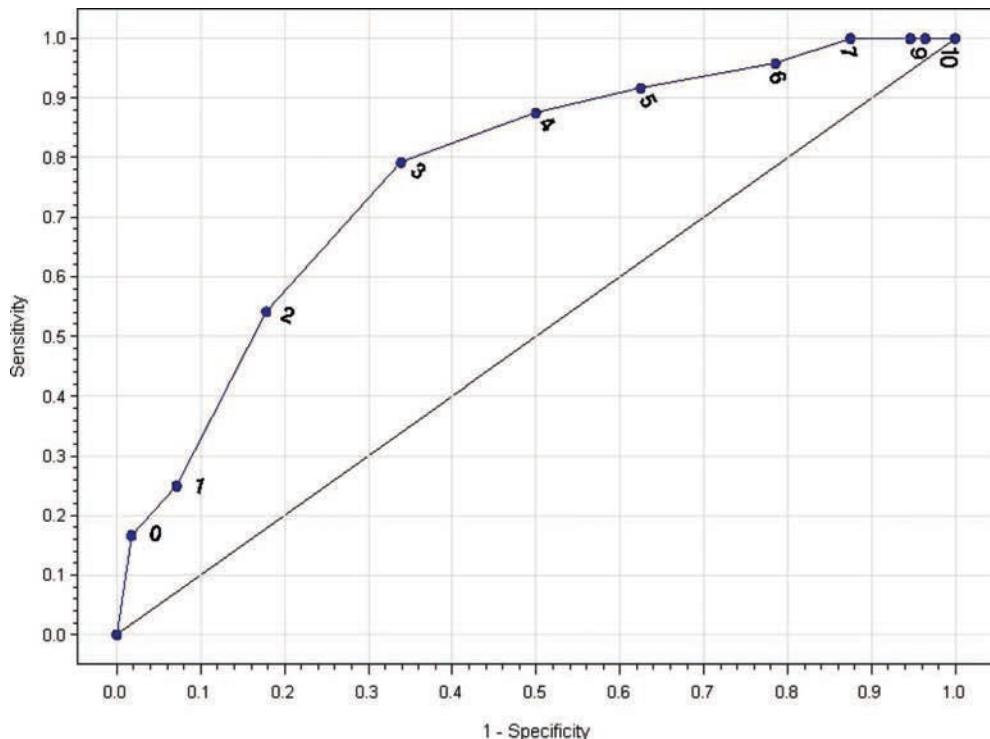


Figure 2. ROC curve for CD based on Bishop score at 37 weeks gestation ( $n = 80$ ). Points on the curve are values of Bishop score. Approximate area under the curve 0.771 (standard error 0.056).

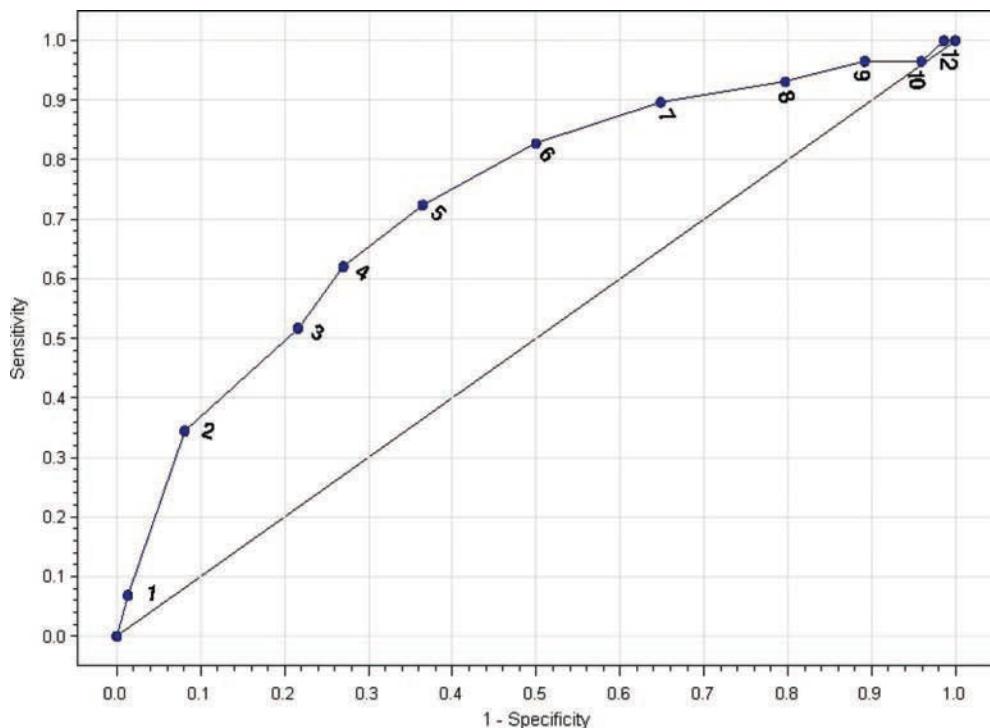


Figure 3. ROC curve for CD based on Bishop score at 39 weeks gestation ( $n = 102$ ). Points on the curve are values of Bishop score. Approximate area under the curve 0.731 (standard error 0.056).

28.6 (53.3) at 40 gestational weeks, and 44.0% (11/25, 95% CI: 26.7–62.9) at 41 gestational weeks.

CD rates for the cohort with Bishop scores  $\leq 5$  was higher than for the cohort with Bishop scores  $> 5$  at each gestational week, based on the gestational age at Bishop score evaluation.

The CD rate was statistically significantly increased by more than two fold at 37, 38, and 39 gestational weeks (Figure 4).

As a binary factor to predict CD, a Bishop score  $\leq 5$  at 37 gestational weeks had an OR of 6.6. This OR decreased with advancing gestational age (Table I). As a continuous factor to

predict CD, each unit increase in the Bishop score decreased the risk of CD. The effect of increased Bishop score varied by gestational week, with the lowest OR (the highest incremental reduction in CD rate with each unit increase in the Bishop score) occurring at 37 gestational weeks (Table II).

The sensitivity, specificity, positive and negative likelihood ratios, and PPVs and NPVs for the Bishop score  $\leq 5$  as a predictor of CD was calculated by gestational week (Table III).

## Discussion

The percentage of nulliparas with a Bishop score  $\leq 5$  decreases after 37 gestational weeks. A Bishop score  $\leq 5$  at any gestational age beyond 37 weeks predicts a higher CD rate compared to patients with a Bishop score  $> 5$  and the OR of eventual CD is highest (OR = 6.6) when the Bishop score is  $\leq 5$  at 37 gestational weeks. At this gestational age, each unit increase in the Bishop score has the greatest impact on reducing the risk of CD (OR 0.60).

Patients with Bishop scores  $\leq 5$  had CD rates that were 2 to 4 times greater than those with Bishop scores  $> 5$ , depending

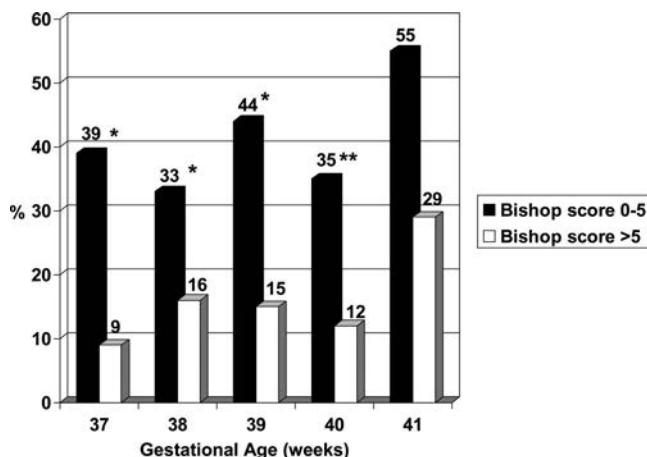


Figure 4. CD rate for enrollees with Bishop scores  $\leq 5$  compared to those with Bishop scores  $> 5$  by gestational age at assessment of score. Sample size by gestational age: 37 weeks,  $N = 80$ ; 38 weeks,  $N = 124$ ; 39 weeks,  $N = 102$ ; 40 weeks,  $N = 57$ ; 41 weeks,  $N = 25$  (\* $p < 0.05$ , \*\* $p < 0.01$ ).

Table I. Using Bishop score  $\leq 5$  as a binary predictor for CD.

Gestational age (weeks)	OR	95% CI
37 ( $n = 80$ )	6.60	1.41 31.0
38 ( $n = 124$ )	2.66	1.05 6.77
39 ( $n = 102$ )	4.56	1.78 11.72
40 ( $n = 57$ )	4.00	1.04 15.44
41 ( $n = 25$ )	3.00	0.57 15.77

Table II. Using Bishop score as a continuous factor to predict CD.

Gestational age (weeks)	OR	95% CI
37 ( $n = 80$ )	0.59	0.44 0.88
38 ( $n = 124$ )	0.85	0.72 1.01
39 ( $n = 102$ )	0.72	0.59 0.87
40 ( $n = 57$ )	0.79	0.61 1.03
41 ( $n = 25$ )	0.86	0.62 1.19

on gestational age. Specifically, the probability of subsequent CD with a Bishop score of  $\leq 5$  at 39 weeks is 44% compared to 15% ( $p < 0.05$ ). This implies an intrinsically higher CD risk in patients with unfavorable Bishop scores at this gestational age, despite expectant management, and may explain why previously published observational studies demonstrate a two fold increased risk of CD in patients induced with an unfavorable cervix [3-9]. These studies compared patients induced with an unfavorable cervix to those who presented in spontaneous labor and thus with a favorable cervix. Our data confirm that patients who are admitted in spontaneous labor have approximately one half the CD rate of those who require induction of labor for a maternal or fetal indication. However, if elective induction were instead compared to expectant management (whose probability of CD is greater than two fold higher among those with an unfavorable cervix compared to those with a favorable cervix) the two groups may show no difference in CD rates since in our study, 35% of patients expectantly managed ultimately required an induction of labor resulting in a 42% CD rate in this group.

It is also possible that by continuing expectant management beyond 39 weeks gestation, the likelihood of CD may actually increase over those patients electively induced since numerous factors which may increase the risk of CD also increase with advancing gestational age such as birth weight, placental insufficiency, oligohydramnios, and preeclampsia.

The strengths of our study include the prospective design and the follow up and evaluation of only those patients at term that had no plan for an indicated induction of labor at 37 weeks gestation. All patients were expectantly managed, meaning that routine obstetric care was provided with weekly visits and delivery planned only for the development of maternal or fetal indications or by 42 weeks gestation. In addition, elective induction of labor is not performed in our institution; therefore, this study is not confounded by this management. Patients only underwent an induction of labor for clinical indications. Adequate numbers of patients also permitted stratification of delivery outcome based on Bishop score and gestational age at examination. Finally, confirmation of gestational age by an ultrasound examination at less than 20 weeks gestation was performed on greater than 90% of patients.

The limitations of this study include the fact that multiple examiners were used to document the Bishop score and interobserver variation in examinations was not performed. However, all providers were experienced upper level residents or attending physicians and a standardized Bishop scoring sheet was used and completed immediately following the examination. Proscribed clinical indications for induction and labor management for each patient studied were not specifically delineated for this study and were left up to the discretion of the attending physician. However, our large

Table III. Sensitivity, specificity, positive (LR+) and negative (LR-) likelihood ratios, and PPV and NPV for Bishop score  $\leq 5$  as a predictor of CD in expectantly managed nulliparous.

Gestational age (weeks)	Sensitivity	Specificity	LR+	LR-	PPV	NPV
37 ( $n = 80$ )	91.7	37.5	1.47	0.22	38.6	91.3
38 ( $n = 124$ )	78.8	41.8	1.35	0.51	32.9	84.4
39 ( $n = 102$ )	72.4	63.5	1.98	0.43	43.8	85.5
40 ( $n = 57$ )	66.7	66.7	2.00	0.50	34.8	88.2
41 ( $n = 25$ )	60.0	66.7	1.80	0.60	54.6	71.4

group practice has no 'private patients' and all clinical decisions are made based on a generally standardized practice pattern with residents managing patients supervised by an in house attending obstetrician. This fact minimizes variation in our clinical practice for both antepartum and intrapartum patients. Another limitation of this study may be that the incidence of CD is related to other factors besides the cervical Bishop score to include body mass index or the presence of gestational diabetes and other medical conditions. However, to our knowledge the strongest predictor of vaginal birth is the cervical condition prior to induction or at presentation in labor, and other maternal or fetal comorbidities are not clearly established modifiers of cervical ripeness [17]. It would have been optimal to know which women progressed to active labor (4 cm) but we unfortunately did not collect this data. Finally, women were enrolled prior to 37 weeks gestation (between 34 and 37 weeks gestation) to provide cervical examination data for the study by this gestational age. Failure to present for care beyond 37 weeks or preterm birth were the most common reasons for exclusion.

This study provides important data for further evaluation of the effect of elective induction of labor compared to expectant management at term. A recent review of this subject and the accompanying clinical commentary highlighted the need for current studies evaluating elective induction of labor at or beyond 39 weeks gestation since there are no recent reports that adequately define the effects of elective induction of nulliparous patients [15,18]. In fact, well designed randomized clinical trials of elective induction of labor versus expectant management at 39 weeks and beyond that are adequately powered to assess important subgroups such as parity and Bishop score at randomization was a specific request of this editorial [18]. Until further data clarifies the effects of elective induction of labor at or beyond 39 weeks gestation, this clinical management should only be offered under an institutional review board approved protocol and elective induction of labor prior to 39 weeks gestation should never be performed or offered if amniotic fluid analysis does not confirm fetal lung maturity because of increased neonatal morbidity.

## Acknowledgment

Presented at the 55th Annual Clinical Meeting/ACOG, May, 2007.

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper, and the views expressed do not represent official policies of the US Federal Government or Department of Defense. Supported by Adeza Biomedical.

## References

1. Friedman EA, Niswander KR, Bayonet Rivera NP, Sachtlenben MR. Relationship of prelabor evaluation in inducibility and the course of labor. *Obstet Gynecol* 1966;28:495-501.
2. American College of Obstetricians and Gynecologists: induction of labor, ACOG Practice Bulletin No. 10. Washington, DC: American College of Obstetricians and Gynecologists; 1999.
3. Yeast JD, Jones A, Poskin M. Induction of labor and the relationship to cesarean delivery: a review of 7001 consecutive inductions. *Am J Obstet Gynecol* 1999;180:628-633.
4. Maslow AS, Sweeny AL. Elective induction of labor as a risk factor for cesarean delivery among low risk women at term. *Obstet Gynecol* 2000;95:917-922.
5. Seyb ST, Berka RJ, Socol ML, Dooley SL. Risk of cesarean delivery with elective induction of labor at term in nulliparous women. *Obstet Gynecol* 1999;94:600-607.
6. Prysak M, Castranova FC. Elective induction versus spontaneous labor: a case control analysis of safety and efficacy. *Obstet Gynecol* 1998;92:47-52.
7. Macer JA, Macer CL, Chan LS. Elective induction versus spontaneous labor: a retrospective study of complications and outcome. *Am J Obstet Gynecol* 1992;166:1690-1697.
8. Johnson DP, Davis NR, Brown AJ. Risk of cesarean delivery after induction at term in nulliparous women with an unfavorable cervix. *Am J Obstet Gynecol* 2003;188:1565-1572.
9. Vahratian A, Zhang J, Troendle JF, Sciscione AC, Hoffman MK. Labor progression and risk of cesarean delivery in electively induced nulliparas. *Obstet Gynecol* 2005;105:698-704.
10. Tylleskar J, Finnstrom O, Leijon I, Hedenskog S, Ryden G. Spontaneous labor and elective induction: a prospective randomized study. I. Effects on mother and fetus. *Acta Obstet Gynecol Scand* 1979;58:513-518.
11. Cole RA, Howie PW, MacNaughton MC. Elective induction of labor: a randomised prospective trial. *Lancet* 1975;1:767-770.
12. Amano K, Saito K, Shoda T, Tani A, Yoshihara H, Nichijima M. Elective induction of labor at 39 weeks of gestation: a prospective randomized trial. *J Obstet Gynaecol Res* 1999;25:33-37.
13. Nielsen PE, Howard BC, Hill CC, Larson PL, Holland RHB, Smith PN. Comparison of elective induction of labor with favorable Bishop scores versus expectant management: a randomized clinical trial. *J Matern Fetal Neonatal Med* 2005;18:59-64.
14. Caughey AB, Sundaram V, Kalmal AJ, Gienger A, Cheng YW, McDonald KM, Shaffer BL, Owens DK, Bravata DM. Systematic review: elective induction of labor versus expectant management of pregnancy. *Ann Intern Med* 2009;151:252-263.
15. Caughey AB, Nicholson JM, Cheng YW, Lyell DJ, Washington AE. Induction of labor and cesarean delivery by gestational age. *Am J Obstet Gynecol* 2006;195:700-705.
16. Caughey AB, Musci TJ. Complications of term pregnancies beyond 37 weeks of gestation. *Obstet Gynecol* 2004;103:57-62.
17. Crane JM. Factors predicting labor induction success: a critical analysis. *Clin Obstet Gynecol* 2006;49(3):573-584.
18. Macones GA. Elective induction of labor: waking the sleeping dogma? *Ann Intern Med* 2009;151:281-282.